Experiment:

20 poker chips. n=

15 trials p=

RED	BLUE

What do you think the true proportion of RED is? Why?

Experiment:

20 poker chips. n=

15 trials p=

RED	BLUE

What do you think the true proportion of RED is? Why?

Experiment:

20 poker chips. n=

15 trials p=

RED	BLUE

What do you think the true proportion of RED is? Why?

CHAPTER 8: Section 1

- What is Inference?

- What is Formal Statistical Inference?

Formal Statistical Inference

The 2 most prominent types of Formal Statistical Inference are….

(1) ____

- this is used for……

(2) ____

- this assesses…..

Notes on Inference

- Inference is based on...

- Inference gives us probabilities...

- When we do inference, we are assuming that the data comes from...

Formal Statistical Inference in general...

-
-
-
-
-

The whole purpose/goal of Statistics...

REVIEW:

Measuring	Statistic	Parameter
mean		
standard deviation		
proportion		

$\hat{p}$ =

What types of problems have we learned that use p (proportions)?

What check did we do when we did these problems?

Why did we do this check? In other words, if the check passed, what did that mean?

What is ?

What is ?

Confidence Intervals

Confidence Interval:

- FORM:

- Example: Presidential Polls:

- Questions:
  - is called an... ___________________
  - What parameter is estimating? _______________
  - The margin of error shows...

    ____
  - The confidence level shows...

    ____

- The two parts of a confidence interval are:

(1) ____

(2) ____

    - This part gives the probability that……

Section 8.1: Confidence Interval for the Population Proportion

- Confidence Intervals are based on…
- Statistic =
- Parameter =
- Check:
- If the check passes, then…
- If is approximately normal then we can use…
- We will be taking a sample size n

FORMULA:

For a confidence interval of the population proportion

GENERIC

(statistic estimate) ± (crit. val)(std. dev. of stat)

means: $\bar{x} \pm (\text{crit val})(\sigma/\sqrt{n})$

SPECIFIC to p

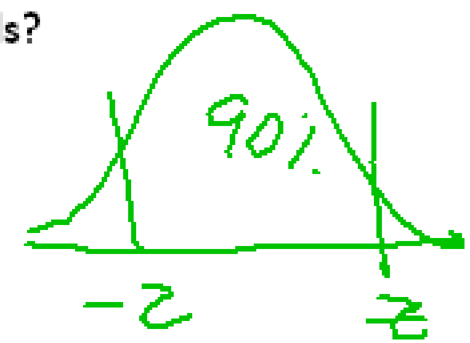
$$\hat{p} \pm z^* \underbrace{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}_{\sigma_{\hat{p}}} = (a, b)$$

stat crit. value formula sheet

- What is Z* defined as?

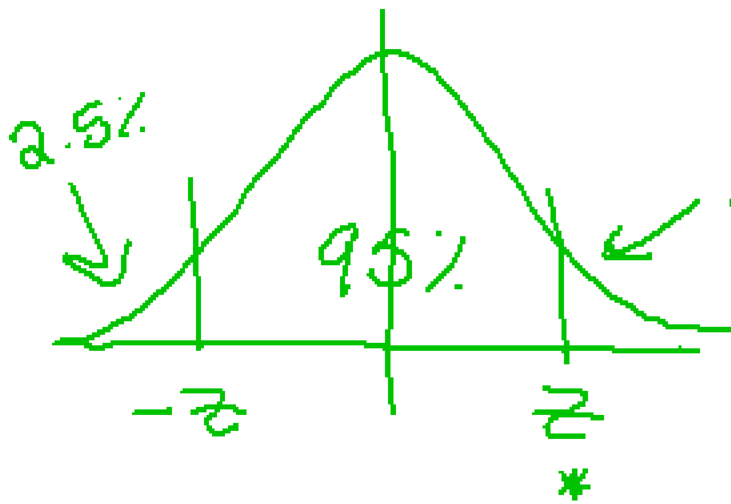
- What is Z* also called? ___________________________

- 3 most common levels of confidence for Confidence Intervals?
  - 90% → 1.645 (z^*)
  - 95%
  - 99%



How can we find Z* for our level of confidence?

- For example, I want the Z* for 9[illegible] 95% confidence level- how can I find it??
  (Look back to your review/warm up for Ch. 6)



$z = \text{invnorm}(.975, 0, 1)$

$=$

3 Common Confidence Levels:

Confidence Level	90%	95%	99%
Z*	1.645	1.96	2.576

But what if we want a different degree of confidence than 90%, 95%, or 99%? How can we find the Z* for those?

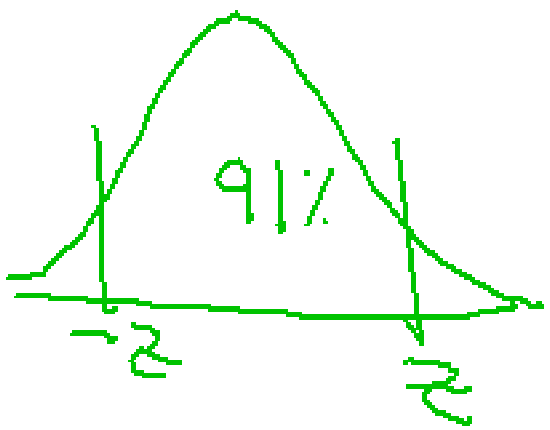
Way #1:

- Look at the tables in the back of the book
- Find Table D (page T-11)
- Down at the bottom of the chart, you will see "Confidence Level C" and a bunch of confidence levels right above it
- Look right above the confidence levels (the %s) and you will see the Z* for those levels of confidence.

Way #2:

- Find the Z* that have that % of data between them using a picture and the calculator function invnorm( ).

Example: Find the Z* for a 91% confidence level:



$z^* = 1.695$

Using the formula for a Confidence Interval, and the information below, complete the following examples.

EXAMPLE #1:

Find a 95% confidence interval when we have $\hat{p} = 0.23$, n = 200.

Check:

$(0.23)(200)$
$(0.77)(200)$ ≥ 10 ✓

46
154 ≥ 10 ✓

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0.23 \pm 1.96 \sqrt{\frac{(0.23)(0.77)}{200}}$$

$$= (0.172, 0.288)$$

EXAMPLE #2:

Alcohol abuse is said to be the #1 problem on college campuses today, and is a leading cause of death for 18-25 year olds. The National Board of Statistics is trying to estimate the percentage of college students that are binge drinkers. They took an SRS of 17096 college students and found 3314 were classified as binge drinkers. (*Note: binge drinking is defined as having 5 or more drinks in a row, 3 or more times in the past 2 week*)

Find a 90% confidence interval for the proportion of college students who are binge drinkers.

$$\boxed{\begin{array}{c} X = 3314 \\ n = 17096 \end{array}}$$

$$\hat{p} = \frac{3314}{17096} = 0.1938$$

Check

3314

13782 ≥ 10 ✓

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0.1938 \pm 1.645 \sqrt{\frac{(0.1938)(1-0.1938)}{17096}}$$

$$(0.189, 0.199)$$

EXAMPLE #3:

Go back to the binge drinking example and find 95% and 99% confidence intervals.

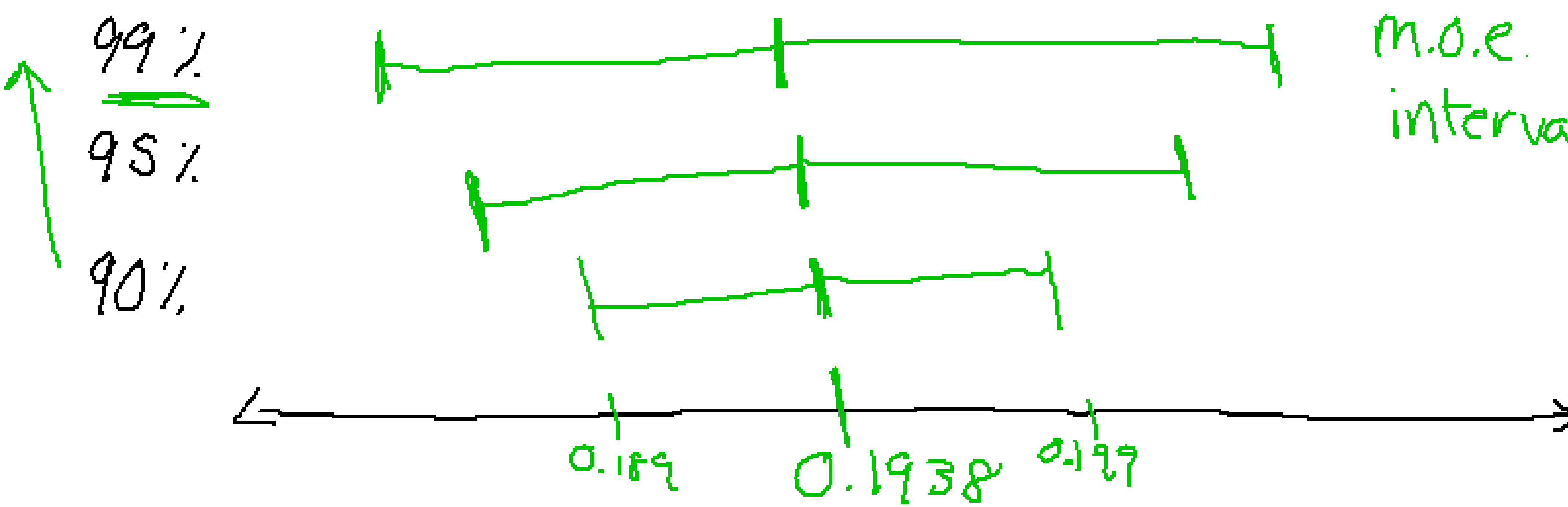
95%: $(0.1879, 0.1998)$

$\hat{p} \pm$ m.o.e

99%: $(0.1861, 0.2016)$

Graph the intervals on the following number line, one on top of the other. What do you notice as the confidence level increases?

as conf level ↑, m.o.e ↑ interval ↑



Interpreting Confidence Intervals

- Confidence intervals have 2 parts:

(1) Interval (a,b)

(2) conf. level (%)

- The interpretation is . . . .

full sentence in context of problem

- It must use both . . . .

interval + conf. level

Form: *(of the sentence interpretation)*

We are ____% confident that the prop. of ________ is btw. ____ and ____.

**Go back to example #2 and interpret your confidence interval.

We are 90% conf. that the prop. of binge drinkers is btw 0.189 and 0.199.

Margin of Error:

- *What part of the formula for a confidence interval is the margin of error?*

$$\boxed{z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

- *Do we want a low or high margin of error?*

- *What 3 things can we do to lower the margin of error?*

z^*

99% 2.576

95% 1.96

* *(1)* ↓ conf. level

* *(2)* ↑ n

(3) $\hat{p}$ further from 0.5

can't

$$\hat{p} = (0.5)(0.5) = 0.25$$

$$\hat{p} = (0.4)(0.6) = 0.24$$

$$\hat{p} \cdot (0.3)(0.7) = 0.21$$

Cautions/Assumptions:

- *A confidence interval is only correct in certain circumstances and with certain assumptions.*

Assumptions: STATE / checks

CHECKS

#'s in

circled / assumed

- $np \geq 10$, $n(1-p) \geq 10$
- SRS
- $pop \geq 10 \cdot n$

$pop \geq 10(50)$ ✓

Cautions:

We can't use our confidence interval if the above assumptions aren't met. Also...

- any other sampling
- biased data

Do the examples on the next page- worksheet 8.1

① assumptions – state & check

② work:
formula & interval
(a, b)

③ interpretation

#1–3 ~ no assumptions

④ State

SRS

np
$n(1-p) \geq 10$

$pop \geq 10n$

Check

- assumed

- $(0.64)(550) \geq 10$ ✓
  $(0.36)(550)$ ✓

- $pop \geq 10(550)$ ✓
  more than 5500 mall patrons

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (0.58728, 0.69272)$$

We are 99% conf. that the prop of mall customers that spent over \$25 is btw. 58.73% and 69.27%.

②

State

SRS

np
$n(1-p) \geq 10$

$pop \geq 10 \cdot n$

Check

circled

$18 \geq 10$
$207 \geq 10$ ✓

$pop \geq 10(225)$ ✓

(more than 2250 machine parts produced)

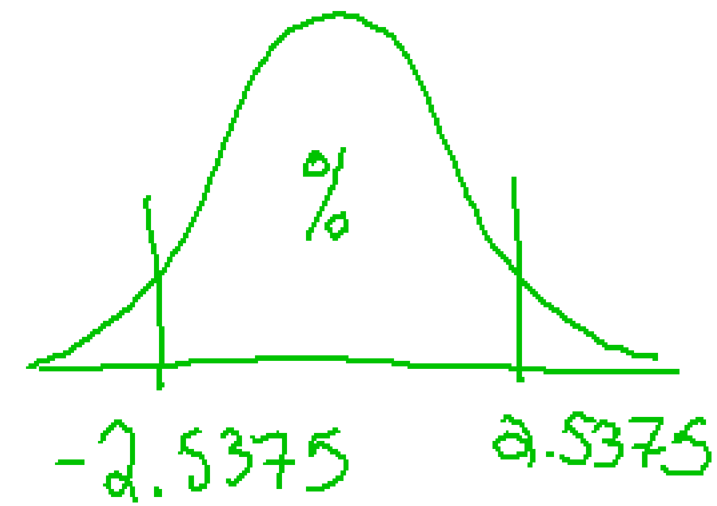
$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (0.04455, 0.11545)$$

We are 95% conf. that the prop. of damaged parts is btw. 0.04455 and 0.11545.

3 a) $n = 1000$

$x = 832$

$$\hat{p} \pm \underbrace{z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}_{m.o.e.}$$



$$0.03 = z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0.03 = z^* \sqrt{\frac{(.832)(1-0.832)}{1000}}$$

$$z^* = 2.5375$$

normcdf(-2.5375, 2.5375, 0, 1)

$= 0.9888$

(b) $83.2\% \pm 3\% = (80.2\%, 86.2\%)$

multiply

pop: 191 mill $(\qquad) \qquad)$